

Confidence Intervals for a Semiparametric Approach to Modeling Nonlinear Relations among Latent Variables

Jolynn Pek, Diane Losardo, and Daniel J. Bauer
University of North Carolina at Chapel Hill

Compared to parametric models, nonparametric and semiparametric approaches to modeling nonlinearity between latent variables have the advantage of recovering global relationships of unknown functional form. Bauer (2005) proposed an indirect application of finite mixtures of structural equation models where latent components are estimated in the service of more flexibly recovering characteristics of the latent aggregate regression function. This article develops and evaluates delta method and parametric bootstrap approaches for obtaining approximate confidence intervals for Bauer's semiparametric approach to modeling latent nonlinear functions. Coverage rates of these approximate point-wise confidence intervals or nonsimultaneous confidence bands are evaluated by Monte Carlo and recommendations for their use are suggested.

Keywords: confidence interval, delta method, parametric bootstrap, structural equation mixture model, nonlinear

Structural equation models are commonly used in the social sciences to describe and test for relationships among latent variables. Although statistical theory for structural equation modeling (SEM) is grounded in linear structural equations, the recognition that nonlinear models might more accurately represent reality has motivated the development of methods for analyzing nonlinear structural equation models (e.g., see Schumacker & Marcoulides, 1998). Historically, parametric approaches to modeling nonlinearity were the first to be proposed (e.g., Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kenny & Judd, 1984; Mooijaart & Bentler, 1986; Ping, 1996) and there remains much continued interest and development in these methods (e.g., Kelava, Moosbrugger, Dimitruk, & Schermelleh-Engel, 2008; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lee & Zhu, 2000; Marsh, Wen, & Hau, 2004). Parametric models explicitly impose some functional form between latent predictor and outcome although the true shape of this function is typically unknown.

More recently, a semiparametric approach to modeling nonlinear relations between latent variables was proposed by Bauer (2005) that involves fitting a structural equation mixture model (SEMM) to the data (Arminger & Stein, 1997; Arminger, Stein, & Wittenberg, 1999; Dolan & van der Maas, 1998; Jedidi, Jagpal, & DeSarbo, 1997a, b; B. O. Muthén, 2001). To date, most interest in SEMMs has centered on direct applications of the model for which the goal is to discern population heterogeneity. In contrast, Bauer's (2005) application of SEMM falls into the indirect category described by Titterton, Smith, and Makov (1985) where mixing components are used as a statistical expedience to more flexibly capture the characteristics of the aggregate population as a whole. In particular, a weighted sum of locally linear within-component relationships is used to estimate the globally nonlinear function. Unlike many of its parametric counterparts, using SEMMs to model nonlinear relationships between latent variables has the advantage of recovering global relationships of unknown functional form without assuming multivariate normally distributed latent variables. Instead, the latent variable distributions are approximated by a mixture of normal distributions that can flexibly assume a wide variety of shapes and the global regression curve is a nonlinear function of the model parameters. As such, no single parameter can be tested to evaluate this nonlinear form as a whole. To enable inference about the function, we propose here two methods for computing approximate point-wise standard errors and confidence intervals or nonsimultaneous confidence bands.

We begin by outlining Bauer's (2005) method of modeling nonlinear latent variable relationships. Next, we describe the delta method and parametric bootstrap resampling approaches to obtaining confidence intervals for the nonlinear function. The coverage rates of these approximate confidence intervals are then evaluated by Monte Carlo for two exemplar functions. We conclude with recommendations for the use of these confidence intervals in practice.

MODEL SPECIFICATION

The SEMM is an extension of the standard structural equation model. To establish notation and facilitate expression of the SEMM, the linear structural equation model is described first. Without loss of generality, the models are presented using one latent predictor η_1 and one latent outcome η_2 . The following developments remain applicable to evaluating bivariate relationships embedded in more complex models.

Linear Structural Equation Model

The measurement model relating continuous observed measures to the latent factors for each individual i is given as

$$\begin{aligned} \mathbf{y}_{1i} &= \mathbf{v}_1 + \lambda_1 \eta_{1i} + \boldsymbol{\varepsilon}_{1i} \\ \mathbf{y}_{2i} &= \mathbf{v}_2 + \lambda_2 \eta_{2i} + \boldsymbol{\varepsilon}_{2i}, \end{aligned} \tag{1}$$

where \mathbf{y}_1 and \mathbf{y}_2 are the vectors of observed variables measuring the latent predictor and outcome, respectively. Intercepts and slopes (loadings) for the regression of the observed

variables onto the latent variables are contained in the vectors \mathbf{v} and $\boldsymbol{\lambda}$, respectively, with subscripts indicating the referent observed variables. The residuals $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ have a joint zero mean vector and covariance matrix $\boldsymbol{\Theta}$. Typically, $\boldsymbol{\Theta}$ is constrained to be diagonal, reflecting the assumption that the observed variables are independent, after conditioning on the latent variables.

The latent variable model for individual i is given by

$$\begin{aligned} \eta_{1i} &= \alpha_1 + \zeta_{1i} \\ \eta_{2i} &= \alpha_2 + \beta_{21}\eta_{1i} + \zeta_{2i}. \end{aligned} \tag{2}$$

The latent predictor η_1 has mean α_1 and variance $VAR(\zeta_{1i}) = \psi_{11}$. Similarly, the latent outcome η_2 has intercept α_2 , slope β_{21} , and residual variance $VAR(\zeta_{2i}) = \psi_{22}$. Further, ζ_1 and ζ_2 are assumed to be independent of each other and of $\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$.

Equations 1 and 2 imply specific mean and covariance structures for the vector of observed variables \mathbf{y} designated as $\boldsymbol{\mu}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, respectively, with $\boldsymbol{\theta}$ representing the vector of model parameters (Bollen, 1989). Assuming that all residuals are normally distributed, the joint marginal probability density function (PDF) is multivariate normal $\phi[\mathbf{y}; \boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})]$, which provides the basis for maximum likelihood (ML) estimation of the parameters. The function $\phi(\cdot)$ denotes the normal PDF.

Semiparametric Nonlinear Effects via Finite Mixture Structural Equation Model

The SEMM assumes that the joint distribution of \mathbf{y} can be approximated by a mixture of K multivariate normal distributions, each parameterized as a linear structural equation model. Each component distribution is usually referred to as a latent class. Let $P(k)$ denote the mixing probability for class $k = 1, 2, \dots, K$ with the constraint that $\sum_{k=1}^K P(k) = 1$. It is assumed that the structural equation model specified for each class has the same form as the measurement model of Equation 1 and this measurement model is constrained to be strictly invariant over latent classes such that the latent variables are equivalently defined for all individuals in the population (Meredith, 1993). Only parameters in Equation 2 must differ across latent classes:

$$\begin{aligned} \eta_{1i} &= \alpha_{1k} + \zeta_{1ki} \\ \eta_{2i} &= \alpha_{2k} + \beta_{21k}\eta_{1i} + \zeta_{2ki}, \end{aligned} \tag{3}$$

as indicated by the subscript k . Within class k , the mean and variance of η_1 are α_{1k} and $VAR(\zeta_{1ki}) = \psi_{11k}$, respectively. The intercept, slope, and residual variance for the linear regression of η_2 onto η_1 are α_{2k} , β_{21k} , and $VAR(\zeta_{2ki}) = \psi_{22k}$. Note that ψ_{11k} and ψ_{22k} can be optionally constrained to be equivalent over classes.

From Equation 3, the expected value of the latent outcome within class k is

$$E_k[\eta_2|\eta_1] = \alpha_{2k} + \beta_{21k}\eta_1 \tag{4}$$

where the relationship between the latent variables is locally linear within each specific class. The potentially nonlinear global relationship between the latent variables is obtained by taking

the expected value across components. The resulting expression aggregates across the locally linear relationships using the mixing probabilities as smoothing weights:

$$E[\eta_2|\eta_1] = \sum_{k=1}^K P(k|\eta_1)E_k[\eta_2|\eta_1] \quad (5)$$

where

$$P(k|\eta_1) = \frac{P(k)\phi_k(\eta_1; \alpha_{1k}, \psi_{11k})}{\sum_{k=1}^K P(k)\phi_k(\eta_1; \alpha_{1k}, \psi_{11k})} \quad (6)$$

is the conditional probability of class membership at a given value of the latent predictor.

In contrast to parametric approaches to modeling nonlinear effects, wherein a single parameter can be tested to evaluate the nonlinear trend, the SEMM approach determines nonlinearity from a function of parameters as in Equation 5. In this context, no single parameter can be tested to permit statistical inferences on the function as a whole. Instead, the nonlinearity can be evaluated by obtaining confidence intervals across the range of the latent predictor.

APPROXIMATE CONFIDENCE INTERVALS

Approximate point-wise confidence intervals or nonsimultaneous confidence bands for the global nonlinear function (Equation 5) can be constructed via the analytical delta method or the empirical bootstrap. With the latter approach, the empirical distribution of the global function can be obtained by nonparametric or parametric forms of the bootstrap. Obtaining reliable SEMM estimates is computationally intensive and prone to problems of nonconvergence, rendering the nonparametric bootstrap cumbersome as SEMMs need to be fit to every bootstrap sample. Hence, we consider only a parametric bootstrapping procedure here where a single SEMM is fit to the sample data. We describe the delta method and parametric bootstrap later in turn.

Delta Method

The delta method provides approximate first and second order moments for nonlinear functions of asymptotically normally distributed estimates. In this instance, the aggregate function (Equation 5) is a nonlinear expression of multiple parameters and can be linearized using a first order Taylor series expansion:

$$E[\eta_2|\eta_1](\hat{\boldsymbol{\theta}}) \approx E[\eta_2|\eta_1](\boldsymbol{\theta}_0) + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)E[\eta_2|\eta_1]'(\boldsymbol{\theta}_0), \quad (7)$$

where $\hat{\boldsymbol{\theta}}$ is the vector of parameter estimates for the K classes, $\boldsymbol{\theta}_0$ is the unknown vector of true parameter values, and $E[\eta_2|\eta_1]'(\boldsymbol{\theta}_0)$ is the vector of first derivatives of Equation 5 taken

with respect to each parameter. The variance of Equation 7 at each specific value of η_1 can be constructed by the following expression:

$$\hat{V}\hat{A}R(E[\eta_2|\eta_1](\hat{\theta})) \approx E[\eta_2|\eta_1]'(\theta_0)|_{\theta_0=\hat{\theta}}^T \hat{V}\hat{A}R(\hat{\theta}) E[\eta_2|\eta_1]'(\theta_0)|_{\theta_0=\hat{\theta}}. \tag{8}$$

Hence, the approximate variance of the aggregate function (Equation 5) is obtained by pre- and post-multiplying the estimated asymptotic covariance matrix of the SEMM ML estimates $\hat{V}\hat{A}R(\hat{\theta})$ by the vector of first derivatives of Equation 7 evaluated at the model implied estimates $\hat{\theta}$. The first derivatives of Equation 7 for K classes are given in the Appendix.

By assuming approximate normality, point-wise confidence bands can be constructed at a desired error rate using the following expression:

$$E[\eta_2|\eta_1](\hat{\theta}) \pm z_{1-\alpha/2} [V\hat{A}R(E[\eta_2|\eta_1](\hat{\theta}))]^{1/2} \tag{9}$$

where $z_{1-\alpha/2}$ is the $\alpha/2$ th quantile of the standard normal distribution. Approximate standard errors associated with a specific value of η_1 are directly obtained by taking the square root of Equation 8. Confidence intervals constructed from Equation 9 are symmetric owing to the symmetry of the normal distribution.

Parametric Bootstrap

The parametric bootstrap is a resampling algorithm such that bootstrap samples are drawn from a parametric model for the data (Efron & Tibshirani, 1993). In the context of modeling latent nonlinearity, the SEMM model defined by Equation 5 is the parametric model of interest. Maximum likelihood estimates of these model parameters can be assumed to follow a multivariate normal distribution:

$$\hat{\theta} \sim N[\theta_0, VAR(\hat{\theta})] \tag{10}$$

where $\hat{\theta}$ remains the vector of ML estimates, θ_0 is the vector of unknown true parameter values, and $VAR(\hat{\theta})$ is the covariance matrix of the estimates. A chosen number of B bootstrap samples are randomly drawn from the parametric estimate of the population defined as:

$$\hat{\theta}_b \sim N[\hat{\theta}, \hat{V}\hat{A}R(\hat{\theta})]. \tag{11}$$

The bootstrapped vector of parameters for sample b , $\hat{\theta}_b$ is assumed to follow a multivariate normal distribution with mean $\hat{\theta}$ and estimated covariance $\hat{V}\hat{A}R(\hat{\theta})$, where $b = 1, 2, \dots, B$. These B bootstrapped values are then used to obtain B estimates of the nonlinear global regression function (Equation 5). These B bootstrapped replicates then serve to estimate the distribution of the global regression. At some desired error rate α , the lower and upper confidence limits for the regression function at a specific value of the latent predictor η_1 are defined as the $[(\alpha/2)100 B]$ th and $[(1 - \alpha/2)100 B]$ th ordered value of the bootstrapped empirical distribution.

One concern with using the parametric bootstrap is that parameter estimates derived from small samples tend to lack precision and variance estimates tend to be nonnormally distributed

and positively skewed. In the context of bootstrapping SEMM model parameters, inadmissible values associated with the K latent predictor variances ψ_{11k} can occur via sampling from Equation 11, especially in small samples. Specifically, a random draw may produce an extremely small or negative value for ψ_{11k} , which leads to improper bootstrapped estimates of the nonlinear global function (Equation 5). Truncation or censoring of the distribution in Equation 11 for ψ_{11k} are two practical solutions to circumvent this problem. Pilot research we conducted indicated that truncated values tend to conform better in distribution to the positive skew inherent in variance estimates compared to censored values. Hence we opted to truncate random draws of ψ_{11k} at an arbitrary value of 0.05; values smaller than 0.05 tended to produce inadmissible bootstrapped estimates. In the Monte Carlo evaluations to follow, we randomly drew 1,500 random samples from Equation 11 and retained the first $B = 1,000$ bootstrapped replicates where $\psi_{11k} > 0.05$.

MONTE CARLO EVALUATIONS

To illustrate the performance of these approximate confidence intervals, data were generated for a symmetric (quadratic) and asymmetric (exponential) nonlinear regression between two latent variables for small ($N = 250$), moderate ($N = 500$), and large ($N = 1,000$) samples using SAS 9.2 (SAS Institute Inc., 2008). Each condition had 1,000 replications. The measurement model of Equation 1 was used to generate data for every case where each latent variable was indicated by six measured variables with means $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{0}_6$ and factor loadings $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \mathbf{1}_6$. The next two sections provide more detail on data generation for the quadratic and exponential regressions, respectively.

Estimation of all models was carried out with *Mplus* 5.21 (L. K. Muthén & Muthén, 2007). Initial starting values were obtained from averaging solutions from 50 replicates estimated with 500 random starts each. From these initial values, 10 additional random perturbations were taken for each replication to reduce the incidence of local solutions. Two-, three-, and four-class models were fit to data from all conditions and results of the best fitting model based on the Akaike Information Criterion (AIC) for the two nonlinear functions are presented. Simulation studies conducted by Bauer, Baldasaro, and Gottfredson (in press) show that selecting the number of classes by AIC results in less bias than the Bayesian Information Criterion (BIC) in recovery of nonlinear functions. All confidence intervals in the following demonstration were constructed to have a nominal 95% coverage rate.

Symmetric Nonlinear Regression

For the symmetric nonlinear function, the measurement model was specified with $\boldsymbol{\varepsilon}_{1i} \sim N(0, 1/3)$ and $\boldsymbol{\varepsilon}_{2i} \sim N(0, .083)$ such that the latent variables explained 75% of the variance of their respective measured variables. The latent variable regression was quadratic of the form

$$\eta_{2i} = 5 - .25\eta_{1i}^2 + \zeta_{2i}, \quad (12)$$

where $\eta_{1i} \sim N(0, 1)$ and $\zeta_{2i} \sim N(0, .25)$ such that 50% of the variance in η_2 was explained by η_1 .

Based on the AIC, the four-class model fit the quadratic data best. Figure 1 depicts the approximate confidence intervals constructed around the estimated aggregate function for a single replication ($N = 500$). The delta method and bootstrapped confidence intervals are similar across much of the range of the latent predictor. The confidence intervals exhibit some slight departure from coincidence in the midrange of the latent predictor, partly due to the asymmetry of the bootstrapped confidence intervals. For both methods, the confidence intervals are wider at the tail ends of the latent predictor, implying less precision in the estimated aggregate nonlinear function at extreme values of η_1 . This behavior is to be expected given that η_1 was simulated from a normal distribution such that there is less information in the tails.

Because the SEMM is applied as an approximation tool, the model is not literally correct and some model error (bias) is to be expected, particularly in regions with little data. Figure 2 presents plots of absolute bias by sample size for 1,000 replications across the range of the latent predictor. Given that there are few datum at the tails of the latent predictor, absolute bias increases with more extreme values of the latent predictor, particularly at low sample size. It can be expected that confidence intervals will similarly perform best within the midrange and

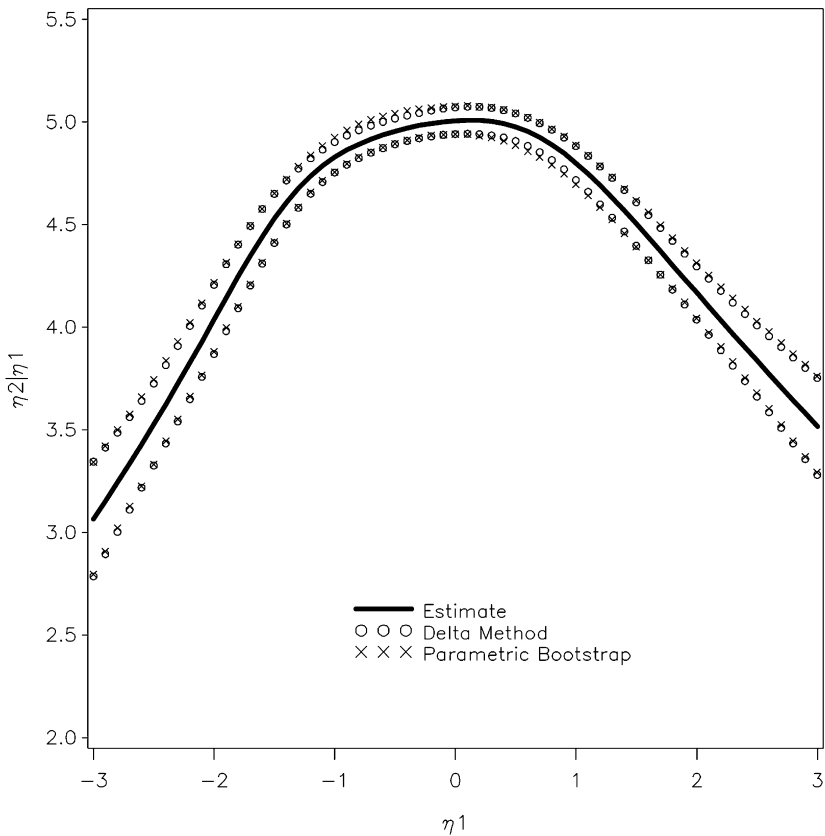


FIGURE 1 Confidence intervals of a single replication for the quadratic function ($N = 500$).

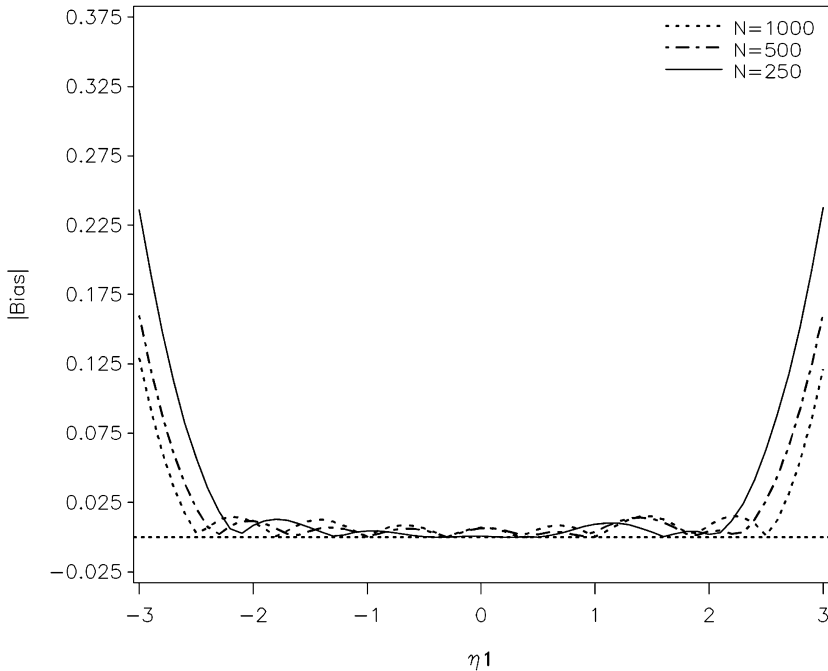


FIGURE 2 Absolute bias of quadratic regression across 1,000 replications.

worst at the tails of the distribution. Indeed, this is precisely the pattern observed in Figure 3, which depicts coverage rates for the quadratic regression across 1,000 replications for the delta method (top) and parametric bootstrap (bottom), respectively. Juxtaposing Figure 2 with Figure 3, coverage rates for both types of confidence intervals tracked absolute bias regardless of sample size. Comparatively, coverage for the bootstrapped confidence intervals mirrored absolute bias more closely than the delta method. Tail-end coverage rates were poorer for the bootstrap method compared to the delta method, whereas the opposite was true for midrange values, owing to the delta method generally having wider intervals.

Coverage rates associated with the delta method confidence intervals were often larger than 95%. Consequently, the precision of the estimates will tend to be underestimated and associated tests will be underpowered. The larger than ideal coverage rates reflect overestimation of the standard errors via the delta method. The upward bias in the delta method standard error estimates is shown in Figure 4, which presents the mean delta method standard errors against the empirical standard error computed from 1,000 replicates for $N = 500$. The delta method approximated standard errors were consistently larger than the empirical standard errors as shown in Figure 4, resulting in the larger than nominal 95% coverage rates observed over much of the range of η_1 in Figure 3.

In contrast, the coverage of the bootstrapped confidence intervals largely maintained the nominal error rate except for extreme values of the latent predictor. Coverage rates for the bootstrapped confidence intervals in Figure 3 varied somewhat by sample size. The moderate

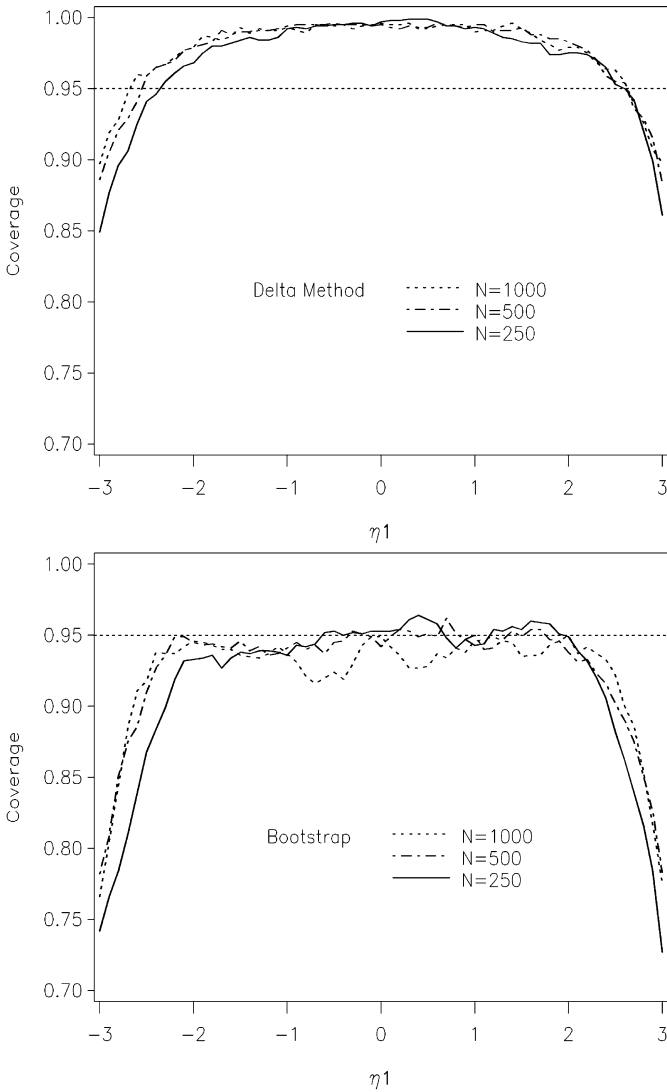


FIGURE 3 Coverage rates of approximate confidence intervals for quadratic regression.

samples had the most ideal coverage and the small samples exhibited the most variability in coverage across the range of η_1 . Contrary to intuition, at certain values of η_1 the bootstrap coverage rates for $N = 1,000$ dip below 95%, especially at the midrange of the latent predictor. One possible explanation for the less than optimal coverage rates with large sample size is potential violation of the parameter (or population) drift assumption (Stroud, 1972; Wald, 1943). This assumption states that the systematic errors attributable to model misspecification are not “too large” relative to sampling errors so as to ensure the existence of asymptotic distributions

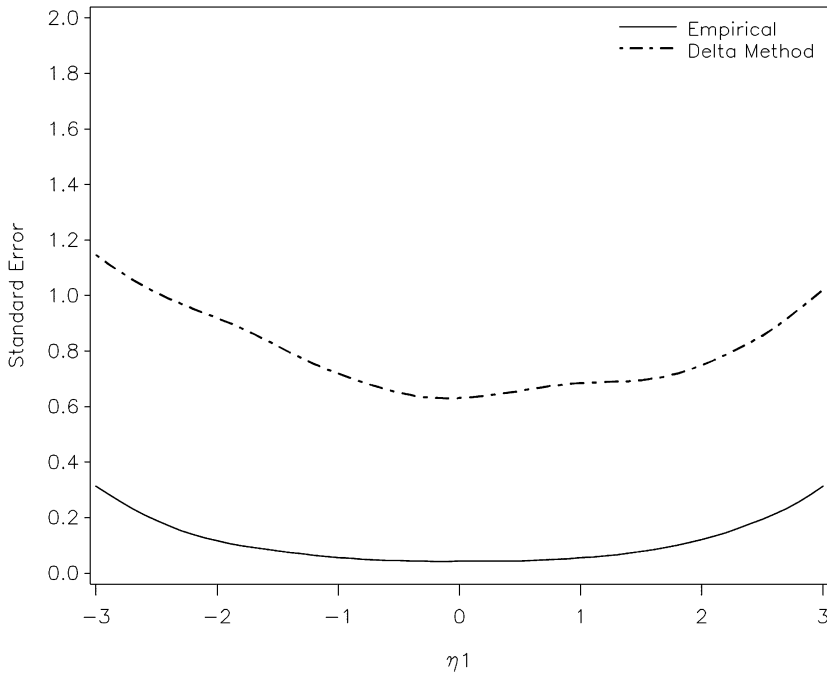


FIGURE 4 Empirical and delta method standard errors for quadratic regression across 1,000 replications ($N = 500$).

(Steiger, Shapiro, & Browne, 1985). By using SEMMs to approximate data generated from a different population function, there will always be some model error. As seen in Figure 2, the model error is relatively constant across different sample sizes, whereas sampling variability necessarily decreases with larger samples. Hence, the ratio of model error against sampling error might be disproportionately large in the context of $N = 1,000$. Violation of the parameter drift assumption may thus have resulted in biased sampling variance estimates.¹ As a consequence, confidence intervals constructed using these suboptimal model estimates tend to unreliably capture the population value.

An alternative explanation centers on the truncation of randomly drawn latent predictor variance estimates ψ_{11k} . Estimates of ψ_{11k} have less precision when obtained from smaller sample sizes, and random draws of these values from Equation 11 in smaller samples tend to result in more inadmissible values. The distribution in Equation 11 thus requires greater truncation (for ψ_{11k}) in smaller samples. Across the 1,000 replications truncation was implemented for 92, 8, and 0 replications for the small, moderate, and large sample conditions, respectively. Therefore, the comparatively greater truncation of the ψ_{11k} distribution at lower sample sizes might have influenced the differential coverage observed in Figure 3.

¹In the context of no model error, bootstrapped nonsimultaneous confidence bands were found to have approximate coverage rates of 95% that improved with increasing sample size.

Asymmetric Nonlinear Regression

For the asymmetric nonlinear function, the errors expressed in Equation 1 were generated as $\epsilon_{1i} \sim N(0, 1)$ and $\epsilon_{2i} \sim N(0, .25)$ such that the latent variables explained 50% of the variance of their respective measured variables. The latent variable model followed the exponential function:

$$\eta_{2i} = 5 + .04[1 - \exp(-1.5\eta_{1i})] + \zeta_{2i} \tag{13}$$

where $\eta_{1i} \sim N(0, 1)$ and $\zeta_{2i} \sim N(0, .12)$.

The three-class model fit the exponential data best with class-invariant parameters based on the AIC. Figure 5 plots the estimated asymmetric aggregate function and corresponding confidence intervals for one replication of moderate sample size ($N = 500$). For this single replication the delta method confidence intervals are slightly wider at the negative tail of η_1 and smaller in the midrange of the latent predictor compared to the bootstrapped confidence intervals.

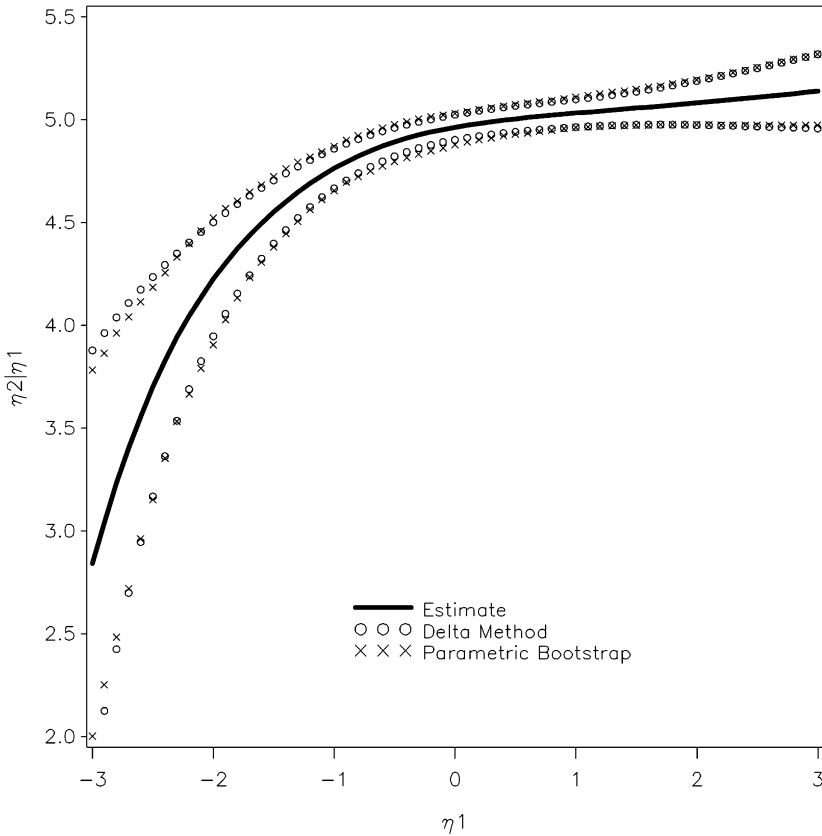


FIGURE 5 Confidence intervals of a single replication for an exponential function ($N = 500$).

A plot of the absolute bias across 1,000 replications by sample size is presented in Figure 6. Similar to the quadratic function, absolute bias tended to increase at the tails of the latent predictor due to data sparseness. In addition, absolute bias for the exponential function was largest toward the negative tail of the latent predictor, where the rate of change in the function is rapid. Given the asymmetric nature of the function, extrapolation of expected values to the right is necessarily more accurate than to the left, producing the asymmetry in Figure 6.

The coverage rates of the confidence intervals are presented in Figure 7. The shape of the coverage curves in Figure 7 reflects absolute bias in Figure 6. Consistent with results from the symmetric function, coverage associated with the bootstrapped confidence intervals exhibited more alignment with absolute bias than the delta method. Additionally, delta method coverage rates were again often larger than 95%. These excessive coverage rates continue to reflect overestimation of the sampling error as shown in Figure 8, which presents the mean delta method standard errors and the empirical standard error computed from 1,000 replicates for $N = 500$. Also consistent with the prior results, sample size did not moderate coverage rates for the delta method under the asymmetric function, but did influence coverage rates for the bootstrapped confidence intervals where the smallest sample size was associated with the most ideal coverage. Again, violation of the parameter drift assumption or truncation of inadmissible draws on the latent predictor variance estimates ψ_{11k} may account for the effect of sample size on bootstrapped confidence interval coverage rates for the asymmetric function. For 1,000

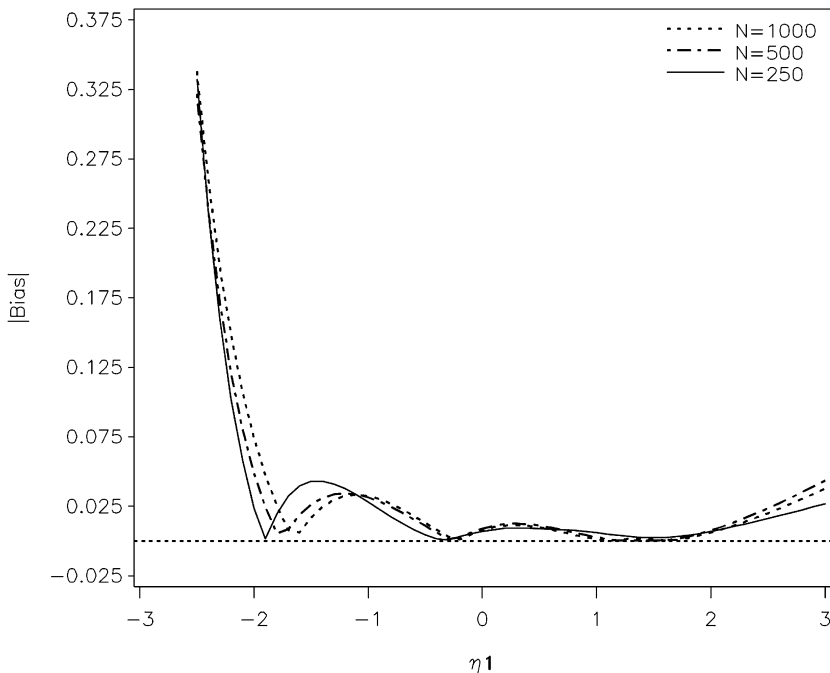


FIGURE 6 Absolute bias of exponential regression across 1,000 replications.

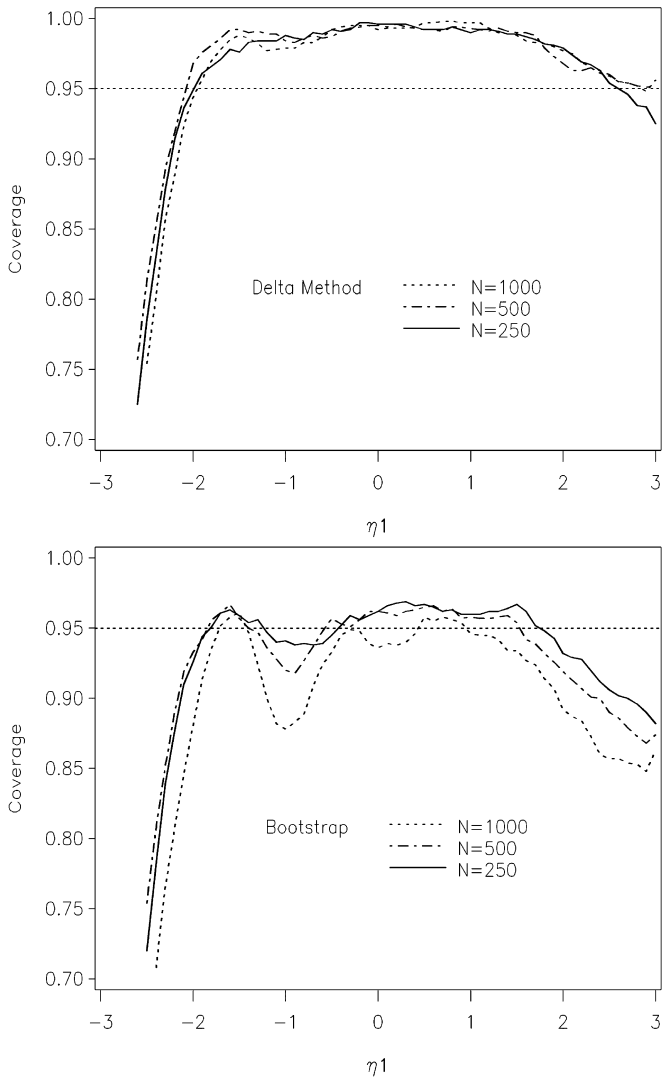


FIGURE 7 Coverage rates of approximate confidence intervals for exponential regression.

replications, truncation was conducted for 300, 58, and 8 replicates for the small, moderate, and large sample conditions, respectively.

CONCLUSIONS AND RECOMMENDATIONS

The precision of nonlinear function estimates obtained from SEMMs can be adequately assessed by approximate point-wise confidence intervals constructed via the delta method or parametric

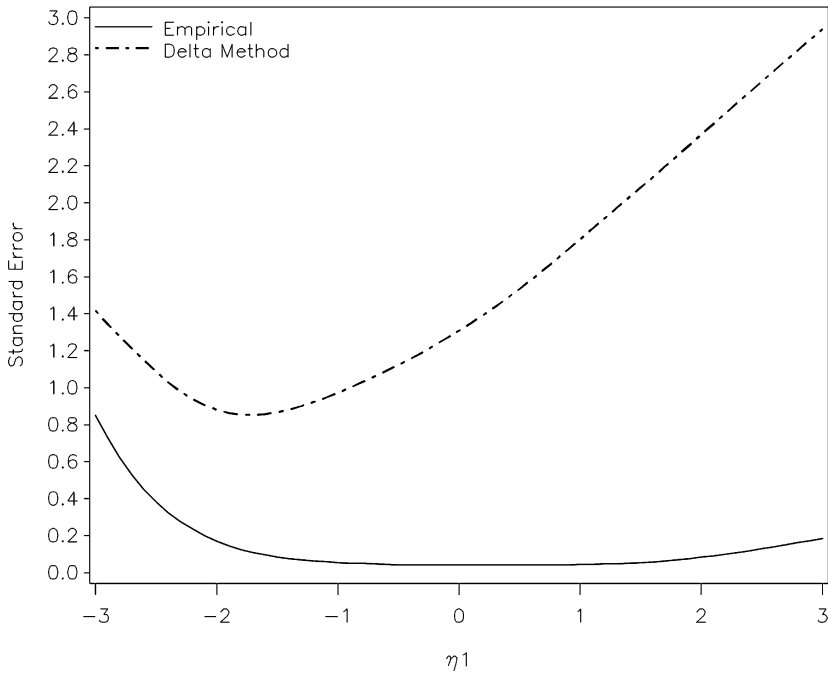


FIGURE 8 Empirical and delta method standard errors for exponential regression across 1,000 replications ($N = 500$).

bootstrap. Confidence intervals obtained from both methods displayed similar characteristics across symmetric and asymmetric functions. Not surprisingly, both methods tended to generate confidence limits with poor coverage at the tails of the latent predictor. This finding is expected given the known bias of the model approximation in regions of sparse data.

Although delta method confidence intervals were somewhat less sensitive to bias, the delta method standard errors consistently overestimated the true sampling error, resulting in coverage rates higher than ideal over much of the range of the function. These confidence intervals then tend to communicate greater estimate imprecision than appropriate, and will inflate the probability of committing a Type II error if used for hypothesis testing. It is possible that the first-order Taylor series provides an insufficiently accurate approximation of the nonlinear function and that a second-order Taylor series approximation might generate more accurate standard errors.

Comparatively, the bootstrapped confidence intervals displayed more ideal coverage rates but tended to track bias closely; coverage rates tended to be liberal at specific ranges of the latent predictor where there was considerable bias. Coverage of the bootstrapped confidence intervals was moderated by sample size, counterintuitively performing best at lower sample sizes. We have suggested that violation of the assumption of parameter drift (Stroud, 1972; Wald, 1943) may be responsible for this finding. Additionally, smaller sample sizes may more often require truncation of inadmissible draws of variance estimates, which may influence coverage

rates. It follows that careful model specification and consideration of sample size serves to guard against potentially suboptimal coverage rates associated with parametric bootstrapped confidence intervals. Choice of the number of latent classes to be modeled becomes key. In general, bias is larger for models with a smaller number of latent classes, and retaining more classes tends to reduce bias and improve bootstrapped confidence interval coverage rates. However, a caveat exists in that taking a larger number of latent classes could result in greater sampling variability, which manifests in wider bootstrapped confidence intervals.

The choice between using the delta method versus the parametric bootstrapped nonsimultaneous confidence bands would depend on analysts' goals in that the two approaches differ in their properties. The delta method confidence intervals were consistently conservative, and tended to cover the population function at the expense of precision. Conversely, the bootstrapped confidence intervals tend to preserve the nominal coverage rate but were sometimes too liberal when bias was present. Although more "tuning" may be required to obtain confidence intervals with desirable properties from the parametric bootstrap, these confidence intervals provide better accuracy in drawing inferences compared to the delta method confidence intervals. In practice, given relatively cheap computational power, the time taken to obtain confidence limits is trivially different between the two approaches. Ultimately, even though performance differed in terms of coverage rates, delta method and bootstrapped confidence intervals are highly similar for any given replication (see Figures 1 and 5).

ACKNOWLEDGMENTS

We would like to thank Robert MacCallum and Michael Browne for helpful discussions regarding the parameter drift assumption. This work was supported by the National Science Foundation Award SES-0716555 to Daniel J. Bauer.

REFERENCES

- Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research*, 26, 148–182.
- Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean and covariance structure models. *Psychometrika*, 64, 475–494.
- Bauer, D. J. (2005). A semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling*, 12, 513–535.
- Bauer, D. J., Baldasaro, R. E., & Gottfredson, N. C. (in press). Diagnostic procedures for detecting nonlinear relationships between latent variables. *Structural Equation Modeling*.
- Bollen, K. A. (1989). *Structural equation models with latent variables*. New York, NY: Wiley.
- Dolan, C. V., & van der Maas, H. L. J. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika*, 63, 227–253.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Jaccard, J., & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, 117, 348–357.
- Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997a). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, 16, 39–59.

- Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997b). STEM: A general finite mixture structural equation model. *Journal of Classification*, *14*, 23–50.
- Jöreskog, K. G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 57–88). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Kelava, A., Moosbrugger, H., Dimitruk, P., & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology*, *4*, 51–66.
- Kenny, D. A., & Judd, C. M. (1984). Estimating the non-linear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201–210.
- Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474.
- Klein, A. G., & Muthén, B. O. (2007). Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*, 647–673.
- Lee, S. Y., & Zhu, H. T. (2000). Statistical analysis of nonlinear structural equation models with continuous and polynomous data. *British Journal of Mathematical and Statistical Psychology*, *53*, 209–232.
- Marsh, H. W., Wen, Z., & Hau, K. T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*, 275–300.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.
- Mooijaart, A., & Bentler, R. (1986). Random polynomial factor analysis. In E. Diday, M. Jambu, L. Lebart, J. Pages, & R. Tomassone (Eds.), *Data analysis and informatics, IV: Proceedings of the fourth international symposium on data analysis and informatics* (pp. 241–250). North Holland, Amsterdam: Elsevier Science.
- Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In A. Sayer & L. Collins (Eds.), *New methods for the analysis of change* (pp. 291–322). Washington, DC: American Psychological Association.
- Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*, 463–469.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Author.
- Ping, R. A. (1996). Latent variable interaction and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin*, *119*, 166–175.
- SAS Institute Inc. (2008). *SAS® 9.2 Enhanced logging facilities*. Cary, NC: Author.
- Schumacker, R. E., & Marcoulides, G. A. (Eds.). (1998). *Interaction and nonlinear effects in structural equation models*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika*, *50*, 253–264.
- Stroud, T. W. (1972). Fixed alternatives and Wald's formulation of the noncentral asymptotic behavior of the likelihood ratio statistic. *Annals of Mathematical Statistics*, *43*, 447–454.
- Titterton, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*. Chichester, UK: Wiley.
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, *54*, 426–482.

APPENDIX

Following B. O. Muthén and Shedden (1999) and B. O. Muthén (2001), the class probabilities are modeled by a multinomial logit regression model for unordered polytomous outcomes where

$$P(k) = \frac{\exp(c_k)}{\sum_k^K \exp(c_k)} \text{ with } c_K = 0. \quad (14)$$

Here, these categorical latent variables c_k are represented as mixture components. Substituting Equation 14 into Equation 6 and simplifying, we have

$$P(k|\eta_1) = \frac{\exp(c_k)\phi_k}{\sum \exp(c_k)\phi_k} \text{ where } \phi_k \text{ denotes } \phi_k(\eta_1; \alpha_{1k}, \psi_{11k}).$$

The derivatives for computing approximate confidence intervals via the delta method follow. Let

$$D = \sum_k^K \exp(c_k)\phi_k.$$

The derivative for the intercept of η_2 for class k is

$$\frac{\partial E[\eta_2|\eta_1]}{\partial \alpha_{2k}} = \frac{\exp(c_k)\phi_k}{D}.$$

The derivative for the slope of η_2 on η_1 for class k is

$$\frac{\partial E[\eta_2|\eta_1]}{\partial \beta_{21k}} = \frac{\exp(c_k)\phi_k \eta_1}{D}.$$

The derivative for the logit for class k is

$$\frac{\partial E[\eta_2|\eta_1]}{\partial c_k} = \frac{\exp(c_k)\phi_k}{D^2} \left\{ \sum_{j \neq k}^K \exp(c_j)\phi_j [\alpha_{2k} - \alpha_{2j}] + [\beta_{21k} - \beta_{21j}]\eta_1 \right\}.$$

The derivative for the mean of η_1 for class k is

$$\frac{\partial E[\eta_2|\eta_1]}{\partial \alpha_{1k}} = \frac{\exp(c_k)\phi'_k}{D^2} \left\{ \sum_{j \neq k}^K \exp(c_j)\phi_j ([\alpha_{2k} - \alpha_{2j}] + [\beta_{21k} - \beta_{21j}]\eta_1) \right\} \text{ where}$$

$$\phi'_k = \frac{\partial E[\eta_2|\eta_1]}{\partial \alpha_{1k}} = \phi_k \left\{ \frac{(\eta_1 - \alpha_{1k})}{\psi_{11k}} \right\}.$$

Finally, the derivative for the variance of η_1 for class k is

$$\frac{\partial E[\eta_2|\eta_1]}{\partial \psi_{11k}} = \frac{\exp(c_k)\tilde{\phi}_k}{D^2} \left\{ \sum_{j \neq k}^K \exp(c_j)\phi_j ([\alpha_{2k} - \alpha_{2j}] + [\beta_{21k} - \beta_{21j}]\eta_1) \right\} \text{ where}$$

$$\tilde{\phi}_k = \frac{\partial E[\eta_2|\eta_1]}{\partial \psi_{11k}} = \frac{\phi_k}{2\psi_{11k}} \left\{ \frac{(\eta_1 - \alpha_{1k})^2}{\psi_{11k}} - 1 \right\}.$$